3 informal

SYNCHRONOUS RESIDUAL TIME STAMP FOR TIMING RECOVERY IN A BROADBAND NETWORK

SU, CL

BACKGROUND OF THE INVENTION

P   This invention relates to timing recovery of a source node service clock

5   frequency at a destination node in a broadband asynchronous mode transfer (ATM) network
where the source and destination nodes receive reference timing signals derived from a
single master clock.

Asynchronous Transfer Mode (ATM) is a packet oriented technology for the
realization of a Broadband Integrated Services Network (BISDN). By using ATM,

10   network resources can be shared among multiple users. Moreover, various services
including voice, video and data can be multiplexed, switched, and transported together
under a universal format. Full integration will likely result in simpler and more efficient
network and service administration and management. However, while conventional
circuit-switching is optimized for real-time, continuous traffic, ATM is more suitable for

15   the transport of bursty traffic such as data. Accommodation of constant bit rate (CBR)
services is, however, an important feature of ATM, both for universal integration and for
compatibility between existing and future networks. In the transport of a CBR signal
through a broadband ATM network, the CBR signal is first segmented into 47-octet units
and then mapped, along with an octet of ATM Type I Adaptation Layer (AAL) overhead,

20   into the 48-octet payload of the cell. The cells are then statistically multiplexed into the
network and routed through the network via ATM switches.

It is essential to the proper delivery of such CBR service traffic in a broadband
network that the clock controlling the destination node buffer be operating at a frequency
precisely matched to that of the service signal input at the source node in order to avoid

25   loss of information due to buffer over- or under-flow. However, unlike the circuit
switched transport of service data wherein the clock frequency at the destination node
may be traced directly back to that of the source node by the regular, periodic arrival of
the CBR traffic, transport in an ATM network inherently results in cell jitter, i.e. the
random delay and aperiodic arrival of cells at a destination node, which essentially

30   destroys the value of cell arrival instances as a means for directly recovering the original
service signal input frequency.

Such cell jitter, generally the result of the multiplexing of transport cells in the
broadband network and the cell queuing delays incurred at the ATM switches in the
network, is substantially unpredictable. Thus, little is known about the cell arrival time

35   beyond the fact that the average cell delay is a constant, assuming that the ATM network
provides sufficient bandwidth to ensure against loss of cells within the network. As a

- 1 -

means for closely approximating the service signal frequency at the destination node, some consideration had previously been given to utilizing a direct extension of circuit switched timing recovery practices which rely entirely upon a buffer fill signal as the basis for recovery of the source timing. However, due to the lack of knowledge of

5   statistics of the cell jitter, this approach would have required a phase-locked loop with very low cut-off frequency (in the order of a few Hz) and would thus have resulted in excessive converging time and degradation of jitter and wander performance.

A number of schemes have been proposed to improve upon such a conventional manner of recovering service timing in the presence of cell jitter, yet none

10   has achieved this end economically and without extensive control systems of notable complexity. Singh et al., for example, in "Adaptive Clock Synchronization Schemes For Real-Time Traffic In Broadband Packet Networks," 8th European Conference on Electrotechnics, Stockholm, Sweden, June 1988, and "Jitter And Clock Recovery For Periodic Traffic In Broadband Packet Networks," IEEE Globecom '88, Florida,

15   December 1988, have proposed algorithms which attempt to more closely estimate cell jitter statistics and derive timing recovery from those indications. These adaptive approaches, suggested to be applicable to both synchronous and non-synchronous networks, rely upon the interaction of increasingly complex algorithms which would require the noted extensive controls for implementation.

20   These prior art schemes described above can be classified as non-synchronous techniques, which are based on the simple fact that the expected value of the network cell jitter is zero and thus rely on phase filtering. Synchronous techniques, on the other hand, utilize the fact that common timing is available at both the transmitter and the receiver. In a synchronous broadband ATM network, such as the Synchronous Optical Network

25   (SONET) prescribed by American National Standard, ANSI T1.105-1988, "Digital Hierarchy Optical Interface Rates and Formats Specification," 10 March 1988, the network source and destination node control clocks are synchronized to the same timing reference. As a result, there is no necessity for relying upon any extraneous phenomenon such as instants of cell arrival to provide a datum base for determining the relative

30   frequencies of those control clocks. The effect of cell jitter caused by multiplexing and switching delays in the network is therefore of little consequence in any procedure for circuit transporting CBR service, which is based, as is the present invention, on an actual synchrony of node timing. Thus being devoid of concern for cell jitter, this process is free to simply determine the difference in frequency between the CBR service signal

35   input at the source node and the source/destination node timing clock(s).

ß U.S. Patent No. 4,961,188 issued on October 2, 1990 to Chi-Leung Lau, co-inventor herein, discloses a synchronous frequency encoding technique (SFET) for clock timing in a broadband network. The SFET takes advantage of the common timing reference at both the source and the receiver. At the source, the asynchronous service

5 clock is compared to the network reference clock. The discrepancy between properly chosen submultiples of the two clocks is measured in units of a preassigned number of slip cycles of network clock. This clock slip information is conveyed via a Frequency Encoded Number (FEN) which is carried in the ATM Adaptation Layer (AAL) overhead. At the receiver, the common network clock and the FEN are used to reconstruct the

10 service clock. This timing recovery process does not rely on any statistics of the cell jitter except that it has a known, bounded amplitude. Therefore, the recovered clock has jitter performance comparable to that of the circuit-switched network.

An alternative proposed approach is known as Time Stamp (TS). In the Time Stamp approach (see, for example, Gonzales et al, "Jitter Reduction in ATM Networks",

15 Proceedings ICC'91, 9.4.1-9.4.6), the network clock is used to drive a multi-bit counter ß(16-bits in the proposal), which is sampled every fixed number of generated cells (e.g., L16). Thus, a fixed number, N, of service clocks cycles is used as the measuring yardstick. The sampled value of the 16-bit counter is the TS that inherently conveys the frequency difference information. Because of the size of the TS (2 octets), it has been proposed that

20 the TS be transmitted via the Convergence Sublayer (CS) overhead. Thus the TS is a 16-bit binary number occurring once every N service clock cycles. Differences in successive TSs represent the quantized values of M, where M is the number of network clock cycles during the fixed TS period. At the receiver, the TS period is reconstructed from the received TSs and the network clock. A free-running 16-bit counter is clocked by

25 the network clock and the output of the counter is compared to the received TSs which are stored in a TS FIFO. A pulse is generated whenever there is a match between the TS and the 16-bit counter. The service clock is recovered by supplying the resultant pulse stream as the reference signal to a multiply-by-N phase locked loop (PLL).

A comparison of the SFET approach and the TS approach reveals advantages

30 and disadvantages for each. In the SFET approach there is a relatively stringent requirement on the derived network clock since it must be slightly larger than the service clock. Advantageously, however, a convergence sublayer is not required to transmit the FEN and only small overhead bandwidth is required to transmit the necessary information. On the other hand, the TS approach is more flexible in that it does not

35 require stringent relationships between the service clock and the network derived clock and can therefore support a range of service bit rates. Disadvantageously, however, a rigid convergence sublayer structure is required to transmit the TS, which adds

complexity and makes inefficient use of the overhead bandwidth.

An object of the present invention is to achieve synchronous timing recovery with an approach that has the advantages of both the SRTS and TS approaches, specifically, the efficiency of SFET and the flexibility of TS.

SUMMARY OF THE INVENTION

5   As described hereinabove, the TS approach requires a large number of bits (16-bits in the example), to represent the number of network clock cycles within a time interval defined by a fixed number (N) of service clock cycles. In accordance with the present invention, the number of bits required to represent the number of network clock

10   cycles within that time interval is substantially reduced. This is possible through the realization that the actual number of network clock cycles, M (where M is not necessarily an integer), deviates from a nominal known number of cycles by a calculable deviation that is a function of N, the frequencies of the network and service clocks, and the tolerance of the service clock. Specifically, therefore, rather than transmitting a digital

15   representation of the quantized actual number of network clock cycles within the interval, only a representation of that number as it exists within a defined window surrounding an expected, or nominal, number of network clock pulses is transmitted from a source node to a destination node in an ATM network. This representation will be referred to hereinafter as the Residual Time Stamp (RTS). By selecting the number of

20   bits, P, so that all 2^P possible different bit patterns uniquely and unambiguously represent the range of possible numbers of network clock cycles within the fixed interval that is defined by N service clock cycles, the destination node can recover the service clock from the common network clock and the received RTS.

At the source node, a free-running P-bit counter counts clock cycles in a clock

25   signal derived from the network clock. The service clock, which is derived from the incoming data signal to be transmitted over the ATM network, is divided by the factor of N to produce a pulse signal having a period (the RTS period) which defines the time interval for measuring the number (modulo 2^P) of derived network clock pulses. At the end of each RTS period, the current count of the free-running P-bit counter is sampled.

30   That sampled value is the RTS, which is transmitted via the adaptation layer. Since the service clock from which the RTS period is defined and the derived network clock are neither synchronized nor integrally related in frequency, the actual number of derived network clock cycles in a RTS period is unlikely to be an integer. Thus, when sampled at the end of each RTS period, the increment in the count of the P-bit counter is a

35   quantized version of the count (modulo 2^P) of pulses in the RTS interval as modified by any accumulated fractional counts from a previous interval.

At the destination node, after the AAL is processed, the successive RTSs are converted into a pulse signal which has periods between pulses defined by the fixed integral numbers of derived network clock pulses that correspond to the conveyed RTS periods. Specifically, a free-running P-bit counter is driven by the derived network

5 clock. A comparator compares this count with a stored received RTS and produces a pulse output upon a match. Since the count of the P-bit counter matches the stored RTS every 2^P derived network clock cycles, comparator output pulses that do not actually represent the end of the RTS period are inhibited by gating circuitry. This gating circuitry includes a second counter that counts the derived network clock cycles

10 occurring since the end of the previous RTS period. When this second counter reaches a count equal to the minimum possible number of derived network clock pulses within an RTS period, the next comparator pulse output produced upon a match between the RTS and the count of the P-bit counter, is gated-through to the output and resets the second counter. The resultant gated through output pulse stream drives a multiply-by-N phase

15 locked loop to recover the service clock .

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 are timing diagrams showing the RTS concept of the present invention;

FIG. 2 is a block diagram showing apparatus, in accordance with the present invention, for generating the RTS at the source node of an ATM network;

20 FIG. 3 is a block diagram showing apparatus, in accordance with the present invention, for reconstructing the service clock at the destination node of an ATM network; and

FIG. 4 are timing diagrams showing the gating function at the apparatus of FIG. 3.

DETAILED DESCRIPTION

The concept of the Residual Time Stamp is described with reference to FIG. 1. In FIG. 1, and in the description hereinafter, the following terminology is used:

f_n ---- network clock frequency, e.g. 155.52 MHz;

f_{nx} ---- derived network clock frequency, $f_{nx} = \dfrac{f_n}{x}$, where x is a rational

30 number;

f_s ---- service clock frequency;

N ---- period of RTS in units of the service clock (f_s) cycles;

- 5 -

T_n ---- the n-th period of the RTS in seconds;

$L \pm \varepsilon$ ---- tolerance of the source clock frequency in parts per million;

M_n (M_{nom}, M_{max}, M_{min}) ---- number of f_{nx} cycles within the $n-th$ (nominal, maximum, minimum) RTS period, which are, in general, non-integers.

5   As can be noted in FIG. 1, during the $n-th$ period, T_n, corresponding to N service clock cycles, there are M_n network derived clock cycles. As aforenoted, since the service clock and the network clock are neither synchronized nor integrally related in frequency, this number of derived network clock cycles is not an integer. Since all practical timing recovery techniques transmit only integer values, the fractional part of

10   M_n must be dealt with. Simple truncation or rounding of the fractional part in each RTS time slot is not permissible, as this would lead to a "random walk" type error accumulation. Rather, it is necessary to accumulate the fractional parts at the transmitter and use the accumulated value to modify the transmitted integer quantity. Since it is most convenient to generate RTS by an asynchronous counter, as will be described

15   hereinafter in conjunction with the description of FIG. 2, a "truncation" operation is natural, reflecting the fact that an asynchronous counter's output does not change until the subsequent input pulse arrives. To formalize these notions, S_n is defined as the truncated value of M_n after accounting for the left over fractional part, d_n, from the (n-1)-th interval, viz.,

20
$$S_n = [M_n + d_n] \tag{1}$$

and

$$d_{n+1} = d_n + M_n - S_n \tag{2}$$

where [a] denotes the largest integer less than or equal to a. Since for accurate clocks, the range of M_n, is very tightly constrained, i. e., $M_{max} - M_{min} = 2y \ll M_n$, the variation

25   in S_n is also much smaller than its magnitude. It follows from Equation (1) that

$$[M_{min} + d_n] \le S_n \le [M_{max} + d_n] \tag{3}$$

Since the maximum and minimum of d_n are 1 and 0 respectively, S_n is bounded by,

$$\left[\frac{M_{\min}}{8}\right] \le S_n \le \left[\frac{M_{\max}}{8}\right] + 1 \tag{4}$$

This implies, that the most significant portion of S_n carries no information and it is necessary to transmit only its least significant portion. This, therefore, is the essential concept of the RTS. The minimum resolution required to represent the residual

5 part of S_n unambiguously is a function of N, the ratio of the network derived frequency to the service frequency, and the service clock tolerance, $\pm\varepsilon$. The maximum deviation, y, between the nominal number of derived network clock pulses in an RTS period, M_{nom}, and the maximum or minimum values of M ($M_{\max}$ or $M_{\min}$) is given by,

$$y = N \times \frac{f_{nx}}{f_s} \times \varepsilon \tag{5}$$

10 where M_{nom} equals $N \times \dfrac{f_{nx}}{f_s}$.

A specific numerical example can be considered for clarity of understanding. As illustrative derived network clock frequency and service clock frequencies could be given by $f_{nx} = 155.52$ MHz (for $x = 1$), and $f_s = 78.16$ MHz (nominal), respectively. A typical RTS sampling period (N) is 3008, which corresponds to a period of 8 cells and a

15 47-octet payload per cell (47 bytes/cell $\times$ 8 bits/byte $\times$ 8 cells per RTS period). Using these numbers, $M_{nom} = 5985.2119$. If it is further reasonable to assume that the service clock tolerance is 200 parts per million, i. e., $\pm 200 \times 10^{-6}$. From equation (5), therefore, y = 1.197, which demonstrates that it is superfluous to transmit the full S_n in each RTS sampling period and transmission of the last few (P) bits of S_n is sufficient. This P-bit

20 sample is the Residual-TS (RTS).

FIG. 2 is a block diagram of the source node of an ATM network showing apparatus for generating and transmitting the RTS. The basic network clock, C, shown at 10, serves as the reference for timing of all nodes of the synchronous network being here considered. This clock, having a frequency f_n, is divided in frequency by a rational

25 factor x by a divider 11 to produce a derived network clock having a frequency f_{nx}. Preferably, x would be an integer value. The dividing factor is chosen so that the P bits available can unambiguously represent the number of derived network clock cycles within an RTS period. In the case where $\dfrac{f_{nx}}{f_s}$ is less than or equal to two, as in the example above, it can be shown that a 3-bit RTS is sufficient.

30 The derived network clock, f_{nx}, drives a P-bit counter, which is continuously counting these derived network clock pulses, modulo 2^P. The service clock, f_s, on lead 113, which is derived from the service data signal (not shown) to be transmitted over the

ATM network, is divided in frequency by N, the desired RTS period in units of f_s cycles, by divide-by N circuit 14. As shown in FIG. 2,, the output of divider 14 is a pulse signal in which T_n is its n-th period. At every T seconds (N source clock cycles) latch 15 samples the current count of counter 12, which is then the P-bit RTS to be transmitted.

5 As aforedescribed, this number represents the residual part of S_n and is all that is necessary to be transmitted to recover the source clock at the destination node of the network.

Each successive RTS is incorporated within the ATM adaptation layer overhead by AAL processor 16. The associated data to be transmitted (not shown) is
10 also processed by processor 16 to form the payload of the cells, which are then assembled by an ATM assembler 17, which adds an ATM header for transmission over the network 18.

B
L With reference again to the previous example, a four-bit counter (P = 4) can be assumed to be used. Since $M_{nom} = 5985.2119$ and 5985.2119 (modulo 16) = 1.2119, a
15 typical RTS output sequence when the source is at nominal frequency will be as follows; PS

Ti B   ....5, 6, 7, 9, 10, 11, 12, 13, 15, 1, 2, ... PS

PS Since the counter 16, in effect, quantizes by truncation, the RTS changes only by integer values. The changes in RTS are such that their average is exactly equal to M_{nom} (modulo
B 2^P). In this example, the changes are either 1 or 2 with the change of 2 occurring either
20 L every 4 or 5 RTSs in such a way that the average interval is 1/0.2119 = 4.7198. In general, successive RTSs are related by PS

Ti
$$RTS_{n+1} = RTS_n + S_n = RTS_n + [d_n + M_n] \ (\text{modulo } 2^P)$$ TM (6) PS

PS In order to guarantee that no information is lost due to the modulo arithmetic, i.e., that the transmitted RTS represents S_n unambiguously, it can be seen from equation (4) that
25 the number of bits used for transmission must satisfy: PS

Ti 8, 9
$$2^P \geq [M_{max}] - [M_{min}] + 2$$ TM (7) PS

PS Thus, in the example above, the number of bits allocated to the RTS must be 3 or greater. It can be noted that the number of bits necessary to unambiguously represent the number of derived network clock cycles within the RTS period is substantially less than the
30 number of bits that would be required to represent the absolute number of clock cycles within the same interval. In the example above, for example, a 13-bit number would be

B

- 8 -

required to represent M_{nom}.

If equation (7) is satisfied, knowledge of M_{nom} in the receiver at the destination node along with the received RTSs can be used to reproduce the service clock from the synchronous network clock. FIG. 3 shows one receiver implementation for reproducing

5   the service clock from the received RTSs. At the receiver the common network clock 10 is available as it was at the transmitter. As in the transmitter, a divider 31 divides the network clock frequency, f_n by the same factor of x as divider 11 in the source node, to produce the same derived network clock signal having a frequency f_{nx} as was used by the transmitter at the source node of FIG. 2.

10   In a structure paralleling the transmitter in FIG. 2, a disassembler 32 processes the ATM headers received from the network 18 and passes the payload to an AAL processor 33. In addition to extracting the transmitted data (not shown), processor 33 extracts the periodic transmitted RTSs, which are sequentially stored in a FIFO 34, which is used to absorb the network cell jitter. The earliest received RTS in FIFO 34 is

15   compared by P-bit comparator 35 with the count of a free running P-bit counter 36, driven by the derived network clock, f_{nx}. Whenever the output of counter 36 matches the current RTS, comparator 35 generates a pulse. Since counter 36 is a modulo 2^P counter, the RTS in FIFO 34 matches the count of counter 36 every 2^P derived network clock pulses, f_{nx}. The output of comparator 35 thus consists of a train of pulses that are

20   separated, except for the first pulse, by 2^P cycles of the derived network clock. In order to select the output pulse of comparator 35 that corresponds to the end of the fixed period of the transmitted service clocks, which is the period per RTS to be recovered, gating circuitry 37 is employed. Gating circuitry 37, which includes a counter 38, a gating signal generator 39, and an AND gate 40, gates only that pulse output of comparator 35

25   produced after counting, from the last gated output pulse, a minimum number, M_l, of derived network clock cycles. This minimum number, M_l, is given by: PS

$$M_l = [M_{nom}] - 2^{(P-1)} \tag{8}$$

PS This ensures that $[M_{max}] - 2^P \le M_l \le [M_{min}]$, and thus the gating pulse is guaranteed to select the correct RTS.

30   P   The gating function is best explained in conjunction with the timing diagrams of FIG. 4. Initially, it can be assumed that gating signal generator 39 is set to keep AND gate 40 open. Comparator 35 compares the first RTS in FIFO 34 with the free-running count of counter 36. When the count of counter 36 matches this first RTS, shown in FIG. 4 as "2", comparator 35 produces a pulse which is gated through AND gate 40. This

35   gated output pulse resets gating signal generator 39 thereupon turning off AND gate 40,

resets the counter of counter 38 to zero, and reads the next stored RTS, "5", in FIFO 34. When counter 36 reaches the count of "5", comparator 35 produces another output pulse. AND gate 40, however, is OFF and remains off until counter 38 counts M_l derived network clock cycles. Therefore, as noted in FIG. 4, all the subsequent matches of the

5   RTS, "5" and the count of counter 36, which occur every 2^P derived network clock cycles, are blocked by AND gate 40. These subsequent pulses are blocked until counter 38 reaches a count of that minimum number of clock cycles that can comprise the fixed interval to be recovered from the RTS. After counting M_l derived network clock cycles, counter 38 generates a pulse which signals gating signal generator 39 to open AND gate

10   40. The next pulse produced by comparator 35 upon the match between the RTS in FIFO 34 and the count of counter 36 is gated through AND gate 40. This pulse, as before, resets counter 38, resets gating signal generator 39, and reads-in the next stored RTS to the output of FIFO 34. The resultant time difference between output pulses of AND gate 40 is the desired fixed time interval, S_n, to be recovered from the transmitted RTSs. As

15   previously defined in equation (1), S_n is the truncated value in the nth interval, after accounting for a left over portion from the (n-1)-th interval, of the actual number of derived network clock cycles within the fixed interval defined by N source clock cycles. As can be noted, S_n modulo (2^P) is equal to the difference of the RTSs associated with the pulses matched by comparator 35 right before and right after the reset. Thus in FIG.

20   4, for the n-th period, this is the difference between "5" and "2", or "3", and for the (n+1)-st period, this is the difference between "9" and "5" or "4". The resultant pulse train at the output of gating circuitry 37 can be seen to duplicate the signal at the source node of the network, which is defined by N service clock cycles, as modified by the quantization effect of the RTSs. This pulse stream is input to a multiply-by N phase

25   locked loop 41 which multiplies the frequency by the factor of N and smooths out the variation of the reproduced periods. The resultant output clock signal, f_r, is the reproduced service timing signal, which can be employed by the circuitry at the destination node.

        The above-described embodiment is illustrative of the principles of the present

30   invention. Other embodiments could be devised by those skilled in the art without departing from the spirit and scope of the present invention.

CM What is claimed is: